Naval Research Laboratory

Washington, DC 20375-5320

NRL/MR/5650--06-9004

Intensity Analysis of Recurrence Plots for the Detection of Deterministic Signals in Noise

Bryan M. Dissinger
G.K. Rohde
R.B. Rhodes, Jr.
F. Bucholtz

Photonics Technology Branch
Optical Sciences Division

J.M. Nichols

Optical Techniques Branch
Optical Sciences Division

December 12, 2006

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
12-12-2006

2. REPORT TYPE
Memorandum Report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Intensity Analysis of Recurrence Plots for the Detection of Deterministic Signals in Noise

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
ONR 61153N

6. AUTHOR(S)
Bryan M. Dissinger, J.M. Nichols, G.K. Rohde, R.B. Rhodes, Jr., and F. Bucholtz

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/MR/5650--06-9004

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
One Liberty Center
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report describes a new method for the detection of deterministic signals in the presence of additive white Gaussian noise. The new method, based on recurrence plots, relies on a sliding-window technique combined with histogram analysis. Performance of the new approach is quantified in terms of receiver-operator characteristic (ROC) curves. The report also provides background information on recurrence plots and ROC curves necessary to understand the new approach.

15. SUBJECT TERMS
Information theory   Signal detection
Statistics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Bryan M. Dissinger
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UL	28	**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified			(202) 767-9498

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

CONTENTS

INTENSITY ANALYSIS OF RECURRENCE PLOTS FOR THE DETECTION OF DETERMINISTIC SIGNALS IN NOISE

Executive Summary

- The detection of deterministic signals buried in noise is a problem of significance for radar warning receivers.

- This report outlines a new method for the detection of deterministic signals based on the use of recurrence plots.

- The new method, called Intensity Analysis, is based on a sliding-window technique applied to the recurrence plot image combined with histogram analysis.

- Performance of the new detection approach, analyzed using ROC curves, shows promise.

- The report also provides background information on recurrence plots and Receiver Operator Characteristic (ROC) curves needed to understand the new approach.

Manuscript approved July 18, 2006.

1. Introduction

1.A. Background

Many present-day radar systems rely on low-probability of intercept (LPI) techniques to avoid detection. An LPI system uses various coding strategies to spread its transmitted energy over a large spectral range. Since the radar receiver "knows" the precise code that was transmitted, target detection can be achieved with high reliability through correlation techniques that yield strong proceeding gain.

From the point of view of the target, the spread-spectrum nature of the radar signal makes detection difficult. In many cases, the power in the radar signal is significantly lower than the thermal (white noise) power of the target's radar-warning receiver (RWR.) For this reason, the detection of deterministic signals in the presence of strong white noise is an important tactical requirement for military platforms.

Detection strategies can be categorized by the amount of *a priori* information assumed. For example, if the practitioner is looking for a sinusoid with known amplitude but unknown phase the quadrature receiver is the optimal choice. If instead the frequency is unknown the practitioner looks at the power in each frequency "bin" of a power spectrum and compares to a threshold. Different scenarios require different detectors. If nothing is known about the incoming signal the problem is much more difficult. The receiver must be designed to look for arbitrary signals with arbitrary (unknown) parameters (e.g. amplitude, phase, frequency, etc.). The optimal receiver in this case will be one that performs well for many types of signals and will likely *not* be optimal for any one signal.

In this work we explore a technique designed to look for the presence of deterministic signals buried in noise that is not based on *a priori* knowledge of the incoming signal. The approach uses Recurrence plots (RPs) to view the time-varying covariance structure of the signal as a binary image. Recurrence plots are in effect time-time distributions and contain similar information to that contained in time-frequency analysis, a frequently used approach in LPI radar applications[1]. RP-based detection strategies have been considered by Zbilut et al. in distinguishing deterministic signals from noise[2,3]. A different approach based on cross-recurrence analysis was also considered for the LPI problem where the incoming signal is known[4], however this approach can be easily shown to be sub-optimal. RPs have also been used in the related problem of de-noising signals[5]. Our approach to recurrence-based detection focuses on the structure of a RP. For Gaussian noise recurrence plots exhibit no structure, that is to say the distribution of points is uniform throughout. If there exists some underlying signal the result is a more structured plot that will consist of dark and light patterns. By forming metrics that quantify these patterns we build RP detectors and demonstrate their performance in detecting signals buried in noise. We discuss the parameters involved in this detection scheme and how they influence the performance of the detector. Results are displayed as Receiver Operating Characteristic (ROC) curves, a commonly used approach to assessing detector performance.

1.B Introduction to recurrence plots

The recurrence plot was originally developed as a graphical means of assessing stationarity in data arising from nonlinear systems[6] and has since been applied to a wide variety of problems (including the LPI applications mentioned above). In short, a recurrence plot quantifies correlations in a time series by tracking the times at which a signal returns to a given state (recurs). Any signal that occupies a finite-dimensional space will, by definition, recur more frequently than infinite-dimensional noise. Thus, recurrence plots and the metrics derived from them provide a reasonable approach to the problem of distinguishing pure noise from signals buried in noise.

The process of constructing a recurrence plot begins with a time-delay embedding of the data. In the analysis of nonlinear dynamical systems, data are often best viewed in state space, which is the space defined by the system's state variables. A system governed by "n" differential equations will therefore be viewed in n-dimensional state space. Nonlinear system identification techniques are then applied to data in this space. In experimental data, however, the practitioner does not typically have access to each of the system's state variables. For example an experimental investigation of a pendulum might allow for the angular position to be measured but not the angular velocity. Fortunately the embedding theorems of Takens[7], and later Sauer[8] et al., demonstrate how delayed copies of a single state variable can be used to generate pseudo-state vectors that preserve certain properties of the "true" underlying dynamical system. The process of embedding allows the practitioner to extract information from a single signal that would otherwise not be available.

We describe the embedding process with a simple example. Consider a 1-dimensional data vector **x** comprising M real numbers

$$x = \left\{ x(1), x(2), \cdots, x(m), \cdots, x(M) \right\}$$

For example, this data may result from the measurement of one parameter such as voltage, frequency, or temperature at discrete times $t = m\Delta t$. (For convenience, Δt is usually understood and it is dropped from the notation so the time is denoted simply by the integer m, sometimes called the time index.) From this single 1-dim data vector of length M, we construct a family of new vectors, each having length $n<M$. The N new vectors X_i are constructed as follows

$$X_i = \left\{ x(i), x(i+L), \ldots, x(i+(n-1)L) \right\}$$

where $1 \le i \le N$, L is the delay time and n is the embedding dimension. Since the largest time index is just M, it must be true that the number of new vectors is

$$N = M - (n-1)L$$

and each of these vectors will be of dimension n. That is, each of the X_i are vectors that live in an n-dimensional space. Provided that n and L are chosen "properly" this representation is topologically equivalent to the underlying n-dimensional system that produced the time series x. There exist a variety of prescriptions for choosing n,L. The optimal delay will produce independent (in the statistical sense) pseudo-state vectors while the optimal dimension will be the smallest one that prevents trajectories in state space from crossing. A thorough discussion of both parameters may be found in Williams[9].

It should be mentioned that the importance of a proper embedding in certain types of recurrence plot analysis is not well understood. The goal in many recurrence-based applications is to analyze patterns in the plot corresponding to correlations at a prescribed length scale (to be discussed shortly) and not necessarily to preserve the underlying dynamics. Additionally there is some research to suggest that estimates of dynamical invariants may be obtained from a recurrence plot without embedding[10]. It may also be difficult to select proper embedding parameters from data that are highly contaminated by noise. In general, we recommend following the proper embedding procedures outlined in references [9].

It is helpful to consider a specific example: M = 1024, n = 4, L = 12. Then N = (1024 - 3*12) = 988 and

$$
\begin{aligned}
X_1 &= \{\, x(1), x(1+12=13), x(1+2*12=25), x(1+3*12=37) \}\\
X_2 &= \{\, x(2), x(2+12=14), x(2+2*12=26), x(2+3*12=38) \}\\
&\vdots\\
X_{988} &= \{\, x(988), x(988+12=1000), x(988+2*12=1012), x(988+3*12=1024) \}
\end{aligned}
$$

Hence, from the original M=1024 element, 1-dimensional data vector we have created a family of N = 988 vectors each having dimension n = 4.

Suppose that we have M =1024 and L = 12 as before but now we choose the embedding dimension to be n = 3. In this case, N = 1024-2*12 = 1000 and

$$
\begin{aligned}
X_1 &= \{\, x(1), x(1+12=13), x(1+2*12=25)) \}\\
X_2 &= \{\, x(2), x(2+12=14), x(2+2*12=26) \}\\
&\vdots\\
X_{1000} &= \{\, x(1000), x(1000+12=1012), x(1000+2*12=1024) \}
\end{aligned}
$$

Next, consider the case M = 1024, L = 1. If n = 1 is chosen, then N = 1024 = M and the X_i are each simply the elements of the original data vector

$$X_1 = \{\, x(1)\}$$
$$X_2 = \{\, x(2)\}$$
$$\vdots$$
$$X_{1024} = \{\, x(1024)\}$$

and the delay time value is inconsequential. On the other hand, if M = 1024, L = 1, and n = 1024, then N = 1024 –1023 = 1 and this original vector X_1 = x is recovered – a pretty uninteresting case.

The real value of this general approach of time-delay embedding arises in the analysis of a nonlinear system that is governed by a number n of differential equations (the dimensionality of the system) and when the available data is the result of the measurement of only a single parameter. By correctly choosing the delay time L and embedding dimension n, certain structures appear in time-delay embedded space that reveal characteristic features of the nonlinear system. Unfortunately, if the dimensionality of the system exceeds n = 3, a full diagram cannot be drawn. The recurrence plot was developed as an attempt to present high dimensionality information in a convenient two-dimensional figure[6].

The method for constructing the RP is now described. Given a 1-dimensional data vector x of length M and a choice of delay time L and embedding dimension n, the N = M – (m-1)L embedded vectors X_i ($1 \le i \le N$) are calculated as described above. Next, a "distance" between any two vectors say, X_i and X_j, is calculated. Denote this distance by $D(i,j) = \left\| X_i - X_j \right\|$. Various definitions of this "distance" exist. The most common is the Euclidean distance between two points in n-dimensional space represented by the vectors X_i and X_j. That is

$$D(i,j) = \left\| X_i - X_j \right\|$$
$$= \left(\left(X_i[1] - X_j[1] \right)^2 + \left(X_i[2] - X_j[2] \right)^2 + \cdots + \left(X_i[N] - X_j[N] \right)^2 \right)^{1/2}$$

where $X_i[k]$ denotes the k[th] element in the vector X_i.

Another possible distance measure is the maximum norm[11]

$$D(i,j) = \left\| X_i - X_j \right\| = \max_k \left(\left(X_i[k] - X_j[k] \right)^2 \right)^{1/2}$$

Once all N^2 distances have been calculated we could construct a surface or contour plot having $N \, x \, N = N^2$ entries corresponding to each of the *D(i,j)* values. However, the standard RP goes one step further and asks whether the distance *D(i,j)* is smaller than some predetermined value ε. If $D(i,j) \le \varepsilon$, then a colored dot (or box) is placed at location *(i,j)* in the RP, otherwise the location is left blank (white). Mathematically the prescription for the RP is written

$$RP(i,j) = \Theta\left(\varepsilon - \left\| X_i - X_j \right\|\right)$$

where Θ is the Heaviside step function

$$\Theta(y) = \begin{cases} 1 & y > 0 \\ 0 & y \le 0 \end{cases}$$

and a colored dot or box is applied wherever $RP(i,j) = 1$ and a white space is applied where $RP(i,j) = 0$.

The following simple example will illustrate the process. Suppose the initial data vector is $x = \{\ 1,2,3,3,0\}$ (hence *M = 5*) and that we choose *L = 1* and *n = 2*. Then *N = 5-1*1=4* and

$$X_1 = \{\ 1,2\}$$
$$X_2 = \{\ 2,3\}$$
$$X_3 = \{\ 3,3\}$$
$$X_4 = \{\ 3,0\}$$

These vectors are plotted in Figure 1.1. Note that in this example all the vectors are unique, $X_i \ne X\ \forall i, j$, but this will not be true in general.

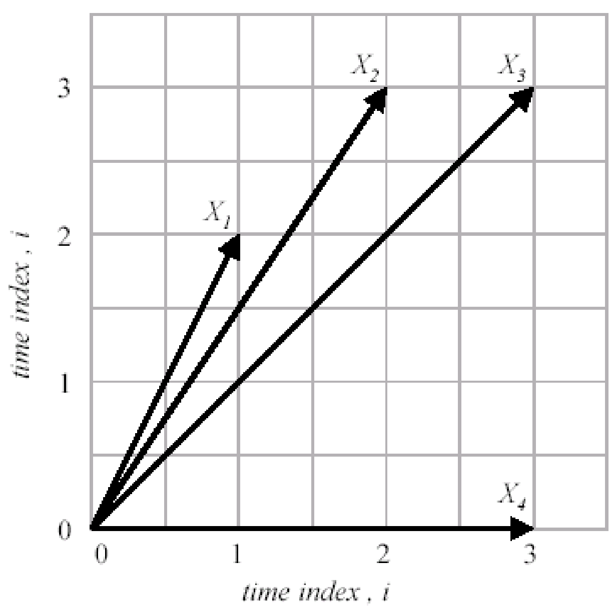

*Figure 1.1: 2-dimensional time-delay embedded vectors created from the data vector x = {1,2,3,3,0} using delay L = 1 and dimension n = 2. Here M = 5 and N = 5-1*1=4.*

Next calculate the distances

$$D(i,j) = \begin{bmatrix} 2\sqrt{2} & \sqrt{10} & 3 & 0 \\ \sqrt{5} & 1 & 0 & 3 \\ \sqrt{2} & 0 & 1 & \sqrt{10} \\ 0 & \sqrt{2} & \sqrt{5} & 2\sqrt{2} \end{bmatrix}$$

Here the lower left entry corresponds to *(i,j) = (1,1)* and the upper right entry corresponds to (4,4). Figure 1.2 shows RPs corresponding to three different values of epsilon: $\varepsilon = 1$, 2, and 3.

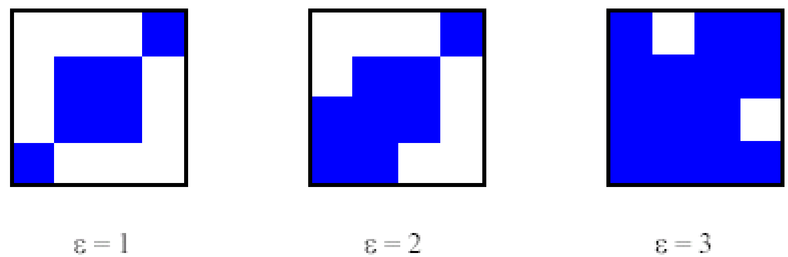

Figure 1.2: Recurrence plots calculated from the time delay embedded vectors of Figure 1.1 for three different values of epsilon.

Since a large amount of possibly high-dimensionality information is forced into a single 2-D graph, the RP must be used with some care. The appearance and the shape of structures in the RP depend on all three plot parameters (L, n and ε) as seen in Figure 1.2.

Finally, we present the contour plot of the *D(i,j)* matrix itself

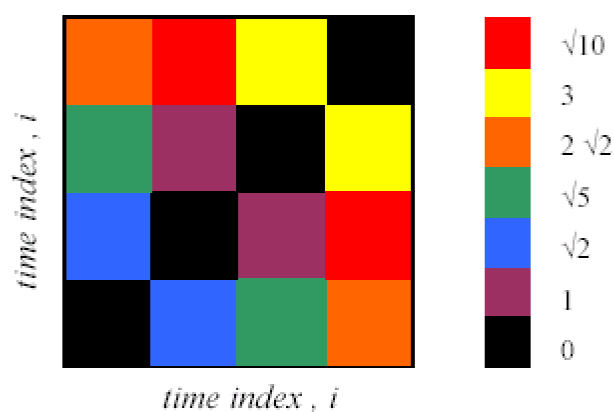

Figure 1.3: Contour plot corresponding to the distance matrix D(i,j).

This report will evaluate the performance of a new detection method based on recurrence plots for both sinusoidal signals and pseudorandom bit sequences (PRBS) contaminated by additive white Gaussian noise (AWGN). In addition to the parameters used in the construction of the RP itself (L, n, ε) we are also concerned with a) the frequency of the signal relative to the sampling rate, and b) the strength of the signal with respect to the strength of the noise. For continuous wave (finite power) signals, the relative strength of signal with respect to noise is the signal-to-noise ratio (snr)

$$snr = \frac{\sigma_{signal}^2}{\sigma_{noise}^2}$$

where σ_{signal}^2 is the variance of the signal and σ_{noise}^2 is the noise variance. Variance is calculated in the usual statistical sense. For data vector $x = \{ x(1), x(2), \cdots, x(M) \}$, the variance is given by

$$\sigma^2 = \frac{1}{M-1} \sum_{j=1}^{M} \left(x(j) - \mu \right)^2$$

and the mean value is $\mu = (1/M) \sum_{j=1}^{M} x(j)$. For example, for a sine wave of amplitude A, $A sin 2\pi ft$, the mean value $\mu = 0$ and the variance $\sigma_{signal}^2 = A^2/2$.

White Gaussian noise is defined by the following characteristics:

1) the mean value of the noise is zero, $\mu = 0$;

2) the envelope of the normalized probability distribution of amplitudes has a Gaussian shape $p(z) = \frac{1}{\sigma_{noise}\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_{noise}^2} \right)$. Hence, the noise variance alone entirely determines the width of the distribution; and

3) the autocorrelation function of the random process is a delta function or, equivalently, the frequency spectrum is flat ("white").

It should be noted that for any finite length data vector, these noise characteristics are met only approximately, not strictly.

In our simulations, the noise is simply added to the signal to give a composite data vector. When the snr is varied for a particular simulation, the noise variance is held constant and the signal level is adjusted to give the desired SNR. For a sinusoidal signal, $A sin 2\pi ft$, the signal-to-noise ratio (linear units) is given by

$$snr = \frac{A^2/2}{\sigma_{noise}^2}$$

and in logarithmic units SNR (dB) = $10\log_{10}(\text{snr})$.   For a PRBS where a zero corresponds to zero and where a one corresponds to A_{PRBS}, the signal variance is $\sigma^2_{signal} = (A_{PRBS}/2)^2$ and the signal-to-noise ratio is

$$snr = \frac{A^2_{PRBS}/4}{\sigma^2_{noise}}.$$

1.C. Commercial software used to produce recurrence plots

The recurrence plots used in this report were all created using Matlab® and the crp toolbox. The crp toolbox provides an easy to use graphical user interface and a command line interface which are appropriately documented. The toolbox is licensed under GNU GPL and is readily available at http://www.agnld.uni-potsdam.de/~marwan/toolbox/ (registration is required.) An online RP generator is also available and can be found at http://www.agnld.uni-potsdam.de/~marwan/rp/rp_www.php.

1.D. Introduction to Receiver Operator Characteristic (ROC) plots

At the most fundamental level, the purpose of any detection system is to inform the user, at any particular time, of the presence or the absence of a signal. Since no detection system is perfect, the user must understand and quantify errors associated with the detection process. Table 1.1 illustrates the four possible detection scenarios: the two correct declarations, namely “declare signal present when the signal is present” and “declare no signal present when no signal is present”; and the two erroneous declarations, namely, “mistakenly declare a signal is present when none is present” that is, a false alarm, and “mistakenly declare no signal when the signal is really there”, that is, a false negative.

Table 1.1: Four possible declaration scenarios for a simple detection system

Detection Scenarios	Declare Signal Present	Declare Signal Absent
Signal Present	True	False
Signal Absent	False	True

A common feature of all detection systems is the notion of a threshold. Suppose the receiver is designed such that, under ideal conditions, when no signal is present exactly zero volts is applied to the detector and, when a signal is present, exactly one volt is applied to the detector. The detector must therefore decide if it has received one volt or zero volts – a task made nontrivial by the fact that any voltage in the real world is contaminated by noise. Rather than asking whether the voltage is exactly zero or exactly one, better results can be obtained by asking if the voltage is less than, or greater than, say 0.5 volts. That is, a threshold is set at 0.5 volts and “one volt” is declared if the voltage is anywhere above the threshold, and "zero volts" is declared if the voltage is anywhere below the 0.5 volt threshold. It is generally true that, for any detection system, the percentage of incorrect or false declarations (Table 1.1) depends on where the threshold is set. In order to quantify this dependence, a particularly useful plot, called the Receiver Operator Characteristic curve, or "ROC" curve, was developed by the radar community.

The ROC curve plots the probability of detection (P_d) on the vertical axis and the probability of false positives, called the probability of false alarm (P_{fa}), on the horizontal axis. Operationally, these probabilities are defined by

$$P_d = \frac{\text{\# times signal declared when signal present}}{\text{total \# times signal present}}$$

$$P_{fa} = \frac{\text{\# times signal declared when no signal is present}}{\text{total \# times signal not actually present}}$$

For example, suppose that, for a particular threshold setting, we conducted 98 trials in which the signal was present for 50 of the trials and absent for 48 trials. Using the same format as for Table 1.1, suppose the results of this experiment are as shown in Table 1.2.

Table 1.2: Example of a detection experiment for n = 98 trials in which the signal was absent for 50 trials and the signal was present for 48 trials.

Detection Scenarios	Declare Signal Present	Declare Signal Absent
Signal Present	45	5
Signal Absent	10	40

In this experiment, then, we have P_d = 45/48 = 0.94 and P_{fa} = 10/50 = 0.20. This result comprises a single point on the ROC curve at (0.20, 0.94). By choosing various threshold levels and conducting the n = 98 trial experiment over and over again, the entire ROC curve can be determined.

2. Recurrence Plots and Signal Detection

In order to use recurrence plots as a method for signal detection a definitive difference between a null plot and signal present plot must be established. The term null plot is a reference to a recurrence plot generated from pure white Gaussian noise (WGN), whereas a signal present plot is generated from a deterministic signal with additive white Gaussian noise. It is important to start with an understanding of how changing the properties of a signal will affect the look of a recurrence plot. There are two major properties of a signal that greatly impact the appearance of a recurrence plot.

The first of these is the signal to noise ratio, where the higher the ratio is, the more defined the structure is within the recurrence plot. As the signal to noise ratio decreases, the more spread out the points of recurrence become, which results in a breakdown of the displayed structure. Figure 2.1 shows a group of recurrence plots with varying levels of signal to noise ratio to illustrate its affect on a recurrence plot.

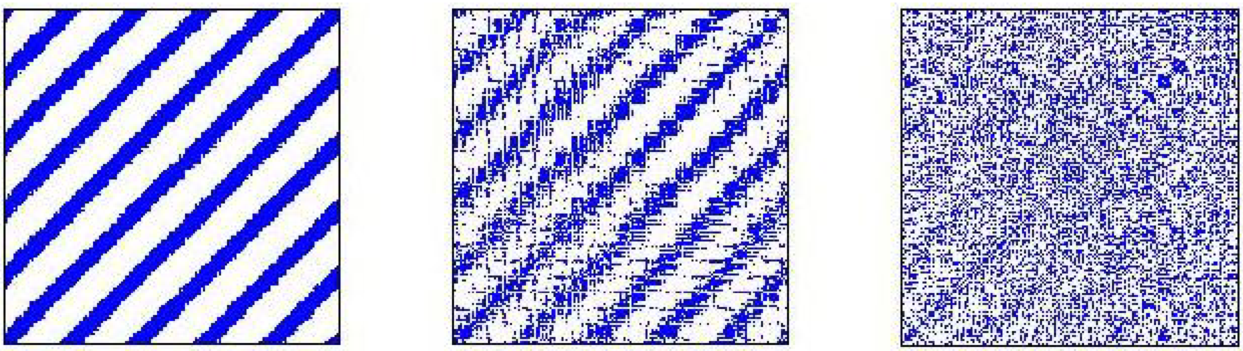

Figure 2.1: RPs created from a 20 Hz harmonic signal with constant frequency and decreasing signal to noise ratio from the left to the right (15dB, 0dB, and –15dB.) The parameters used in creating the recurrence plots were dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors.

All recurrence plots displayed in this report were created from time-series vectors consisting of 1024 data points (M = 1024.) In order to display these recurrence plots in an appreciable way, each recurrence plots shown is only a fraction of the original plot. Each plot originally contained 988 by 988 pixels, whereas the shown portion is an enlarged view of the first 250 by 250 pixels.

Along with signal to noise ratio, the ratio of signal frequency to sampling frequency greatly impacts the quality of the generated recurrence plot. When the signal frequency is low, relative to the sampling frequency, the lines of structure within the recurrence plot tend to be more spaced out. As the amount of oversampling is lessened, the lines of structure are drawn closer together. The thicknesses of the lines are highly dependent on the epsilon used to create the RPs. An illustration of this can be seen if Figure 2.2.

Figures 2.1 and 2.2 clearly show how signal to noise ratio and oversampling independently affect recurrence plots, whereas Figure 2.3 shows their combined affects.

The left column is identical to that of Figure 2.2, whereas the middle and the right columns were created from the same signals with lower signal to noise ratios.

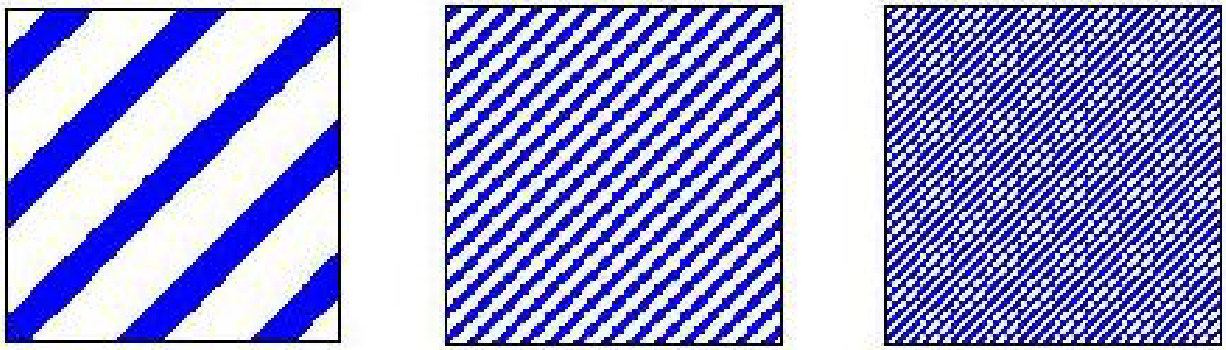

Figure 2.2: RPs created from a harmonic signal with a constant sampling rate (1 kHz) and an increasing frequency from the left to the right (10 Hz, 50 Hz, and 100 Hz.) The parameters used in creating the recurrence plots were dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors.

Notice how the structure within the recurrence plots becomes less visible as you go down the right column even though their signal to noise ratios are identical. The reason for this is because the spreading of the points of recurrence has a much greater impact when the lines of structure are thinner and spaced closer together, based upon the epsilon and oversampling, to begin with. The level of noise needed to blur the recurrence plot, making it indistinguishable from a null plot, is much less in this case. Therefore, the less oversampled a signal is, and the lower the signal to noise ratio is, the quicker the structure within the recurrence plot deteriorates, along with the chances of signal detection.

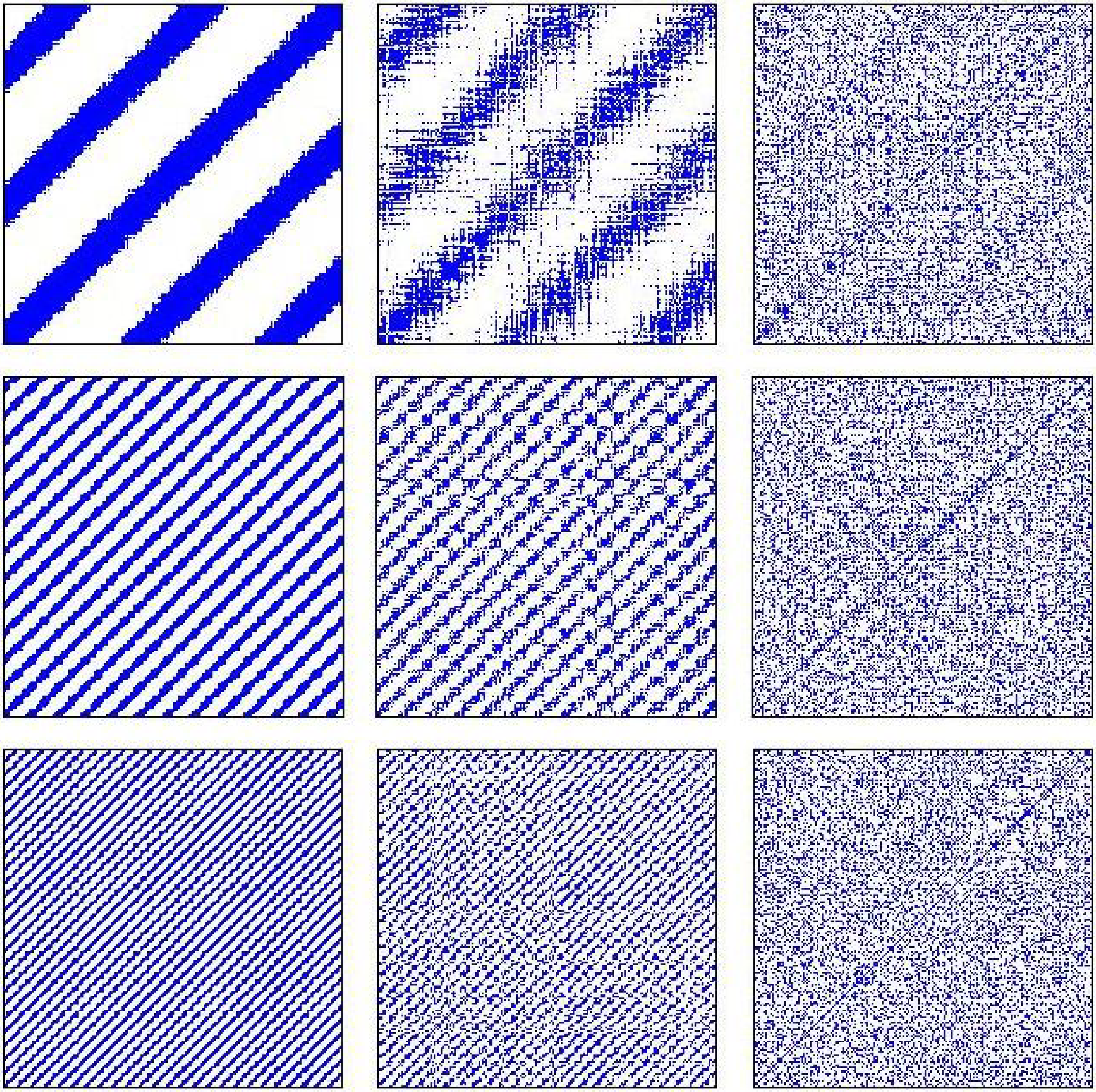

Figure 2.3: RPs created from a harmonic signal with constant sampling rate (1 kHz) and increasing frequency from the top to the bottom (10 Hz, 50 Hz, and 100 Hz) and decreasing signal to noise ratio from the left to the right (15dB, 0dB, and –15dB.) The parameters used in creating the recurrence plots were dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors.

3. Intensity Analysis

When signals are visibly detectable in a recurrence plot it is due to the varying intensity, or the change in the spatial distribution, of the recurrence points. Even though these variations are what the human eye naturally picks out, having a processor identify these can be a little more challenging when the plot becomes blurred. Intensity Analysis is a proposed technique designed to characterize the localized variations in the points of recurrence within a recurrence plot in order to detect the presence of a deterministic signal embedded in noise.

Intensity Analysis is accomplished by sliding an overlapping window, one pixel at a time (i.e. 2-D convolution,) over the entire plot and summing up the number of recurrence points contained within the window. The results can then be analyzed and a metric can be created to determine whether or not a deterministic signal was present. Figure 3.1 shows an example of how Intensity Analysis is performed. The recurrence plot on the left is an enlarged view of a portion of a recurrence plot, which was created from a harmonic signal that was sampled twenty times faster than its frequency, and has a signal to noise ratio of –10 dB.

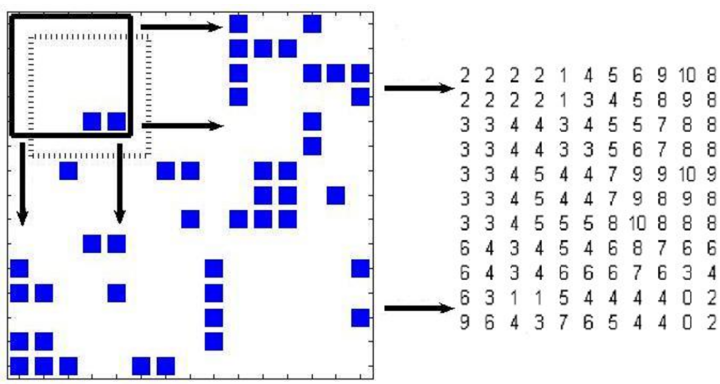

Figure 3.1: Example of Intensity Analysis using a 5x5 window. The plot on the left is an enlarged view of a small portion of a recurrence plot and the numbers on the right are the results from Intensity Analysis.

An appealing way to view the results from Intensity Analysis is to construct a histogram from the results to show the distribution of the points of recurrence. Figures 3.2 shows some results created from signals with various levels of AWGN. Viewing the plots across the rows, from the top to the bottom, the signal to noise ratios are 10 dB, 0 dB, -5 dB, -10 dB, -15 dB, and pure WGN. The x-axis in the histograms represents the resultant value for the number of points of recurrence within a given window. For example, with a window size of five by five, the x-axis will have *w* values from zero to twenty-five. The y-axis corresponds to the number of times that resultant value was found throughout the recurrence plot.

These shapes of the histograms should not be surprising because it follows what has all ready been observed. The histograms with the larger tail sections and lower central region, corresponding to a higher signal to noise ratio, were a direct result from the well-defined lines of structure within the recurrence plots. Conversely, the histograms with the smaller tail sections and larger peaking central region, corresponding to the weaker signal strength, were due to the spreading of the points of recurrence within the recurrence plots. Hence, the more evenly distributed the points of recurrence become (i.e. the further the signal is buried in noise,) the larger the central peak becomes, and the lower the tails become.

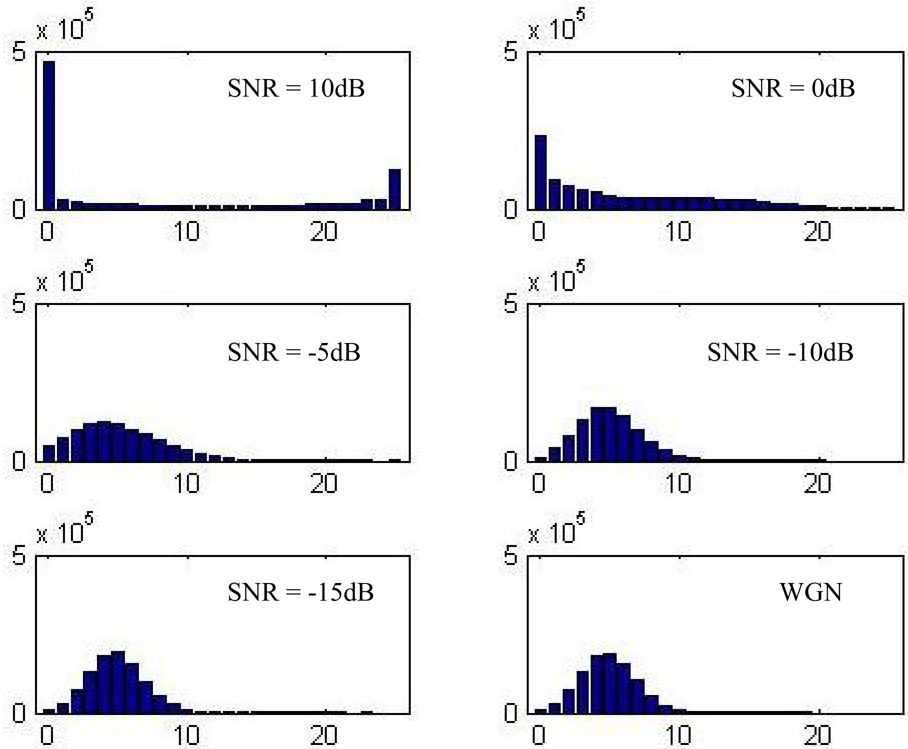

Figure 3.2: Histogram of results from Intensity Analysis using a 5 by 5 window. Each result was taken from a 20 Hz harmonic signal, sampled at 1 kHz, with decreasing signal to noise ratio from the left to the right and from the top to the bottom. The bottom right result is from a pure white Gaussian noise signal. The corresponding recurrence plots that these histograms were created from can be seen in Figure 3.3.

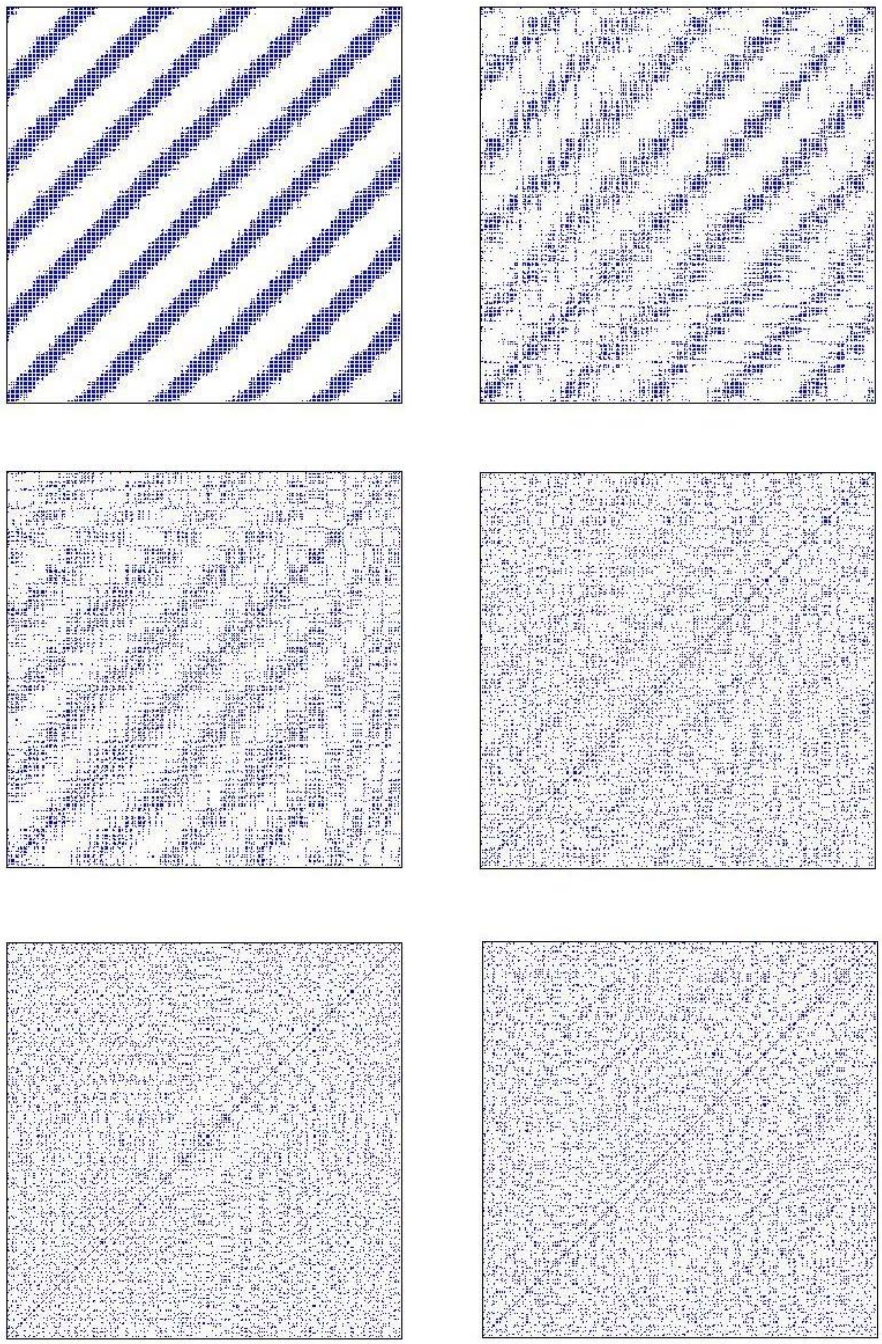

Figure 3.3: Recurrence Plots used to create the histograms shown in Figure 3.2. The RPs were created from a 20 Hz harmonic signal sampled at 1kHz. The signal to noise ratios are decreasing from the left to the right and from the top to the bottom (10dB, 0dB, -5dB, -10dB, -15db, and white Gaussian noise.) The parameters used in creating the recurrence plots were dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors.

4. Results

4.A. Intensity Analysis

Looking at the histograms from Figure 3.2, as the signal to noise ratio was decreased, the shapes of the histograms became very similar to that of pure white Gaussian noise. Figure 4.1 has the results overlaid to show how similar they are. Three key areas of the results are also enlarged to show the trends at these points for the differing levels of signal to noise ratio.

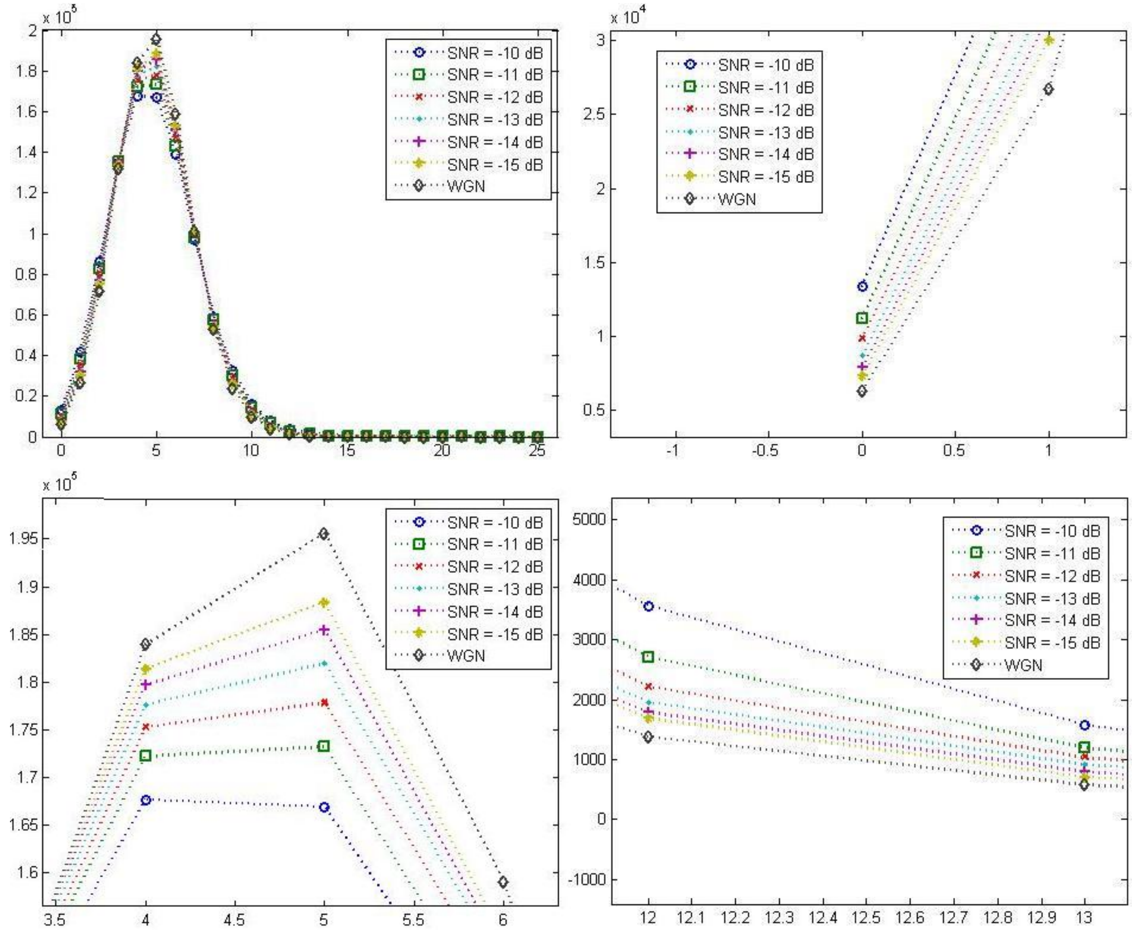

Figure 4.1: These plots show how signal to noise ratio impacts Intensity Analysis of recurrence plots, highlighting the key regions of the results.

The results show that as the signal to noise ratio is decreased the tails of the histogram become smaller while the peak becomes larger. More importantly, as the signal to noise ratio is decreased, the points converge with those of pure white Gaussian noise. It should also be noted that these signals used the same signal and noise vectors in order to better demonstrate the trends (i.e. the signal vector was multiplied by some constant while the noise vector remained constant in order to obtain the desired SNR.)

Figure 3.2 and 4.1 clearly illustrate the effects of signal to noise ratio on Intensity Analysis, but does not expound on the effects of oversampling. Figure 4.2 compares two signals with different levels of oversampling, and a relatively low signal to noise ratio, to a pure Gaussian noise signal. Notice how much closer to the Gaussian noise signal the less oversampled signal is. These results reinforce what was seen earlier where a recurrence plot deteriorates at quicker rate when noise is added for a signal that is less oversampled.

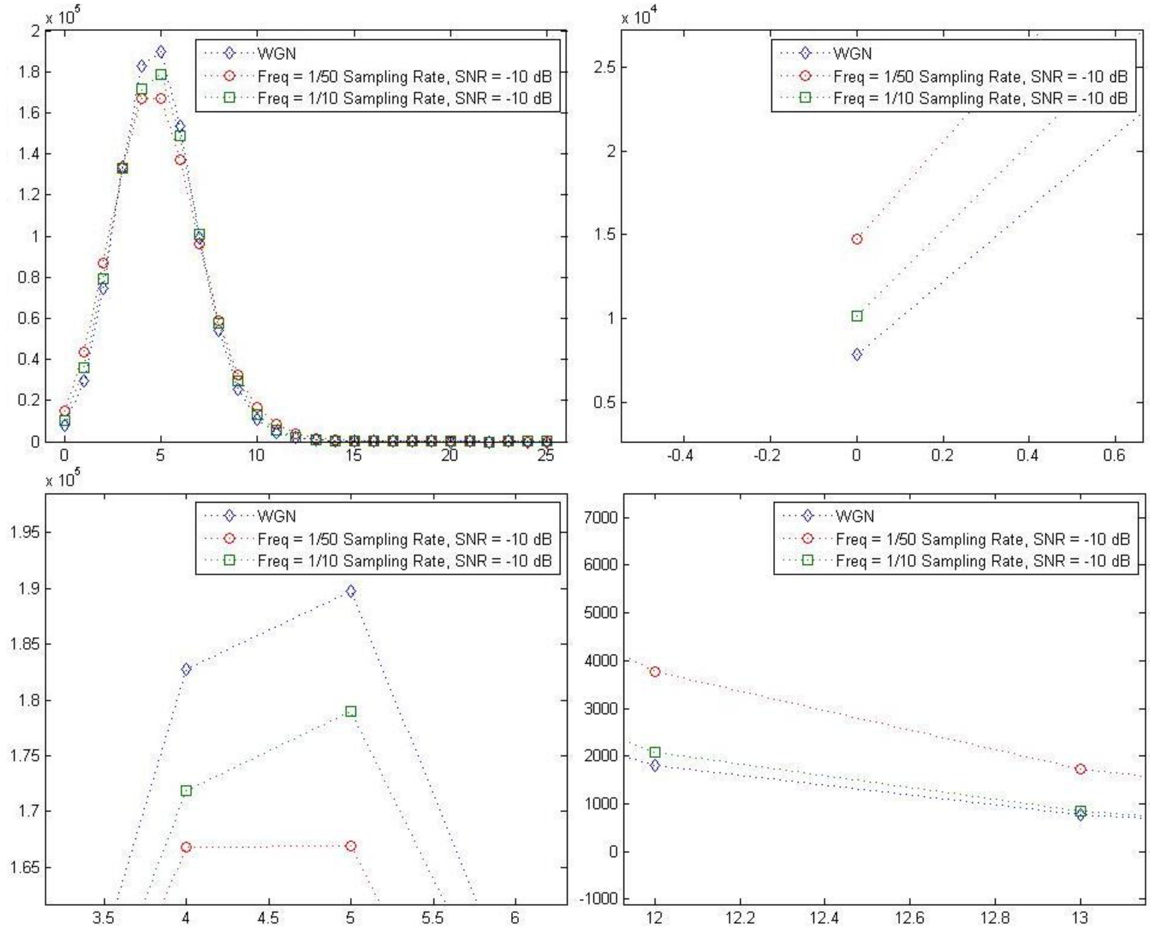

Figure 4.2: These plots show how the ratio of signal frequency to sampling frequency impacts Intensity Analysis of recurrence plots, highlighting the key regions of the results.

Since recurrence plots will vary when generated from signals containing white Gaussian noise, as seen in Figure 4.3, it can only be expected that the results from Intensity Analysis will follow. The amount of this variation is an important factor in being able to detect the presence of a deterministic signal embedded in noise. Figure 4.4 illustrates the amount of variation can be expected from Intensity Analysis results.

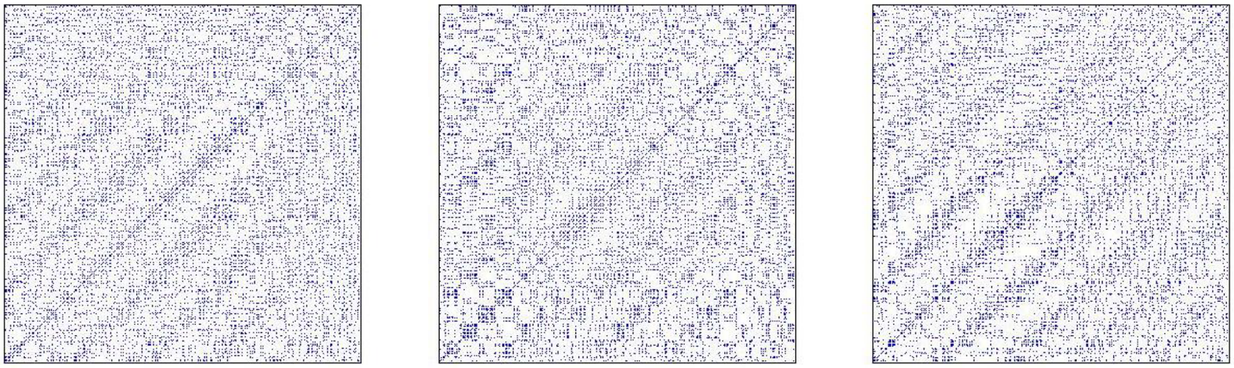

Figure 4.3: These plots show the variation in recurrence plots generated from signals with the same signal to noise ratio (-10 dB) and amount of oversampling (sampling frequency is 50 times greater than the harmonic frequency.) The parameters used in creating the recurrence plots were dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors.

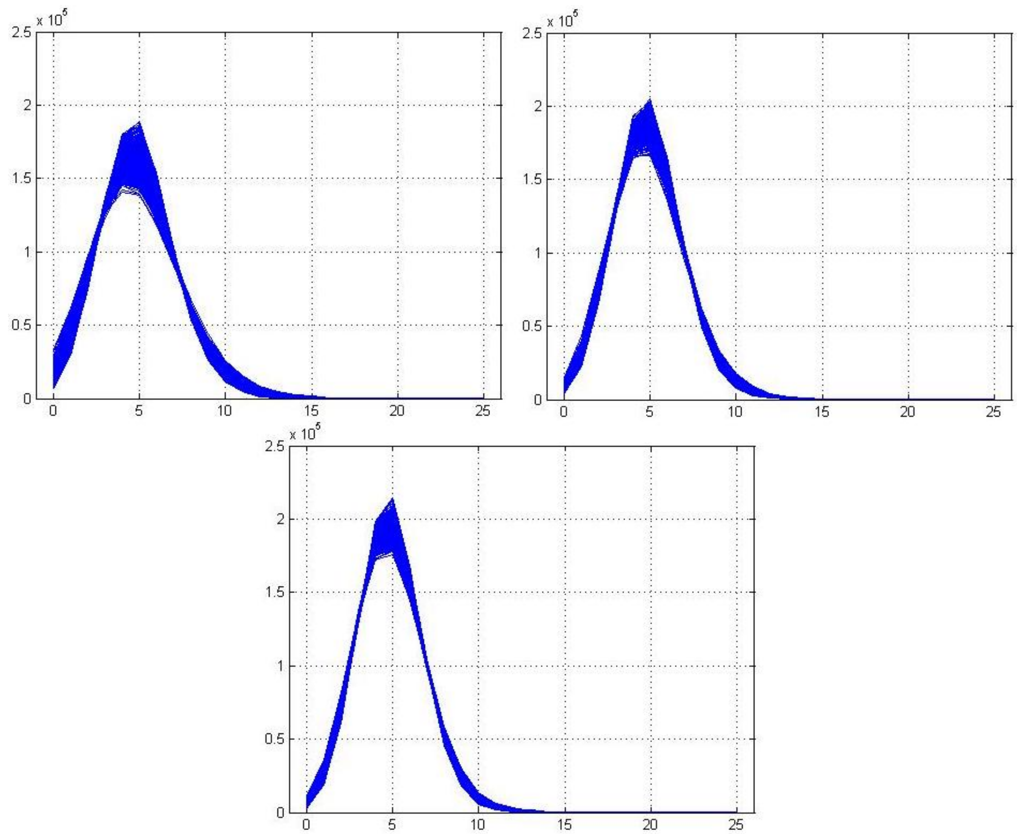

Figure 4.4: These plots show the variation that is seen in Intensity Analysis results. Each plot was generated from an ensemble of 1000 sample runs. The top left plot was created from harmonic signals with a frequency of 1/50th the sampling rate and a signal to noise ratio = -10 dB; the top right plot was created from harmonic signals with a frequency of 1/10th the sampling rate and a signal to noise ratio of –10 dB; the bottom plot was created from pure white Gaussian noise signals.

4.B. ROC Curves

To test the feasibility of using Intensity Analysis of recurrence plots to determine the presence of a deterministic signal embedded in noise, a ROC curve, as described in Section 1.B, can be used. Note that drawing a diagonal from the (0,0) to the (1,1) coordinates would represent a fifty-fifty chance (pure guessing) of correctly detecting the presence of a deterministic signal embedded in noise.

There are many tests that could be constructed and applied to the results of Intensity Analysis to determine whether or not a deterministic signal is present. From the Intensity Analysis results displayed in Figures 4.1 and 4.2, the zero term, central hump or peak term, and variance seem like logical test points. As previously discussed, the zero term is expected to be lower on average for pure GN signals whereas the peak term is expected to be higher. The variance is expected to be larger as well for pure GN signals since the histograms for these signals tend to be more peaked and drawn in.

Both harmonic signals and PRBS signals were created with varying SNRs and varying frequencies and compared to that of pure white Gaussian noise signals. The following figures are the resultant ROC curves from the previously mentioned tests.

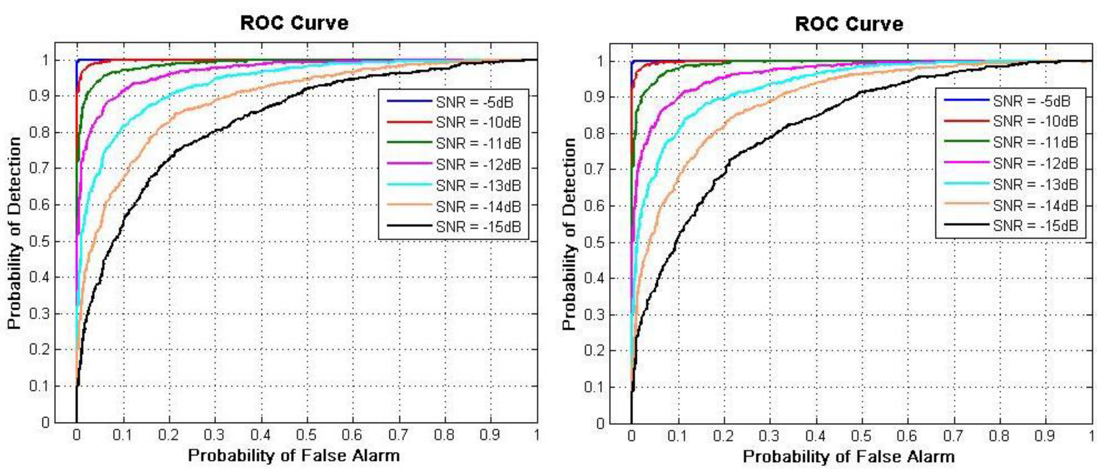

Figure 4.5: Signal detection results of varying SNR levels for the zero term. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. Here the zero term was thresholded to determine the presence, or absence, of a signal.

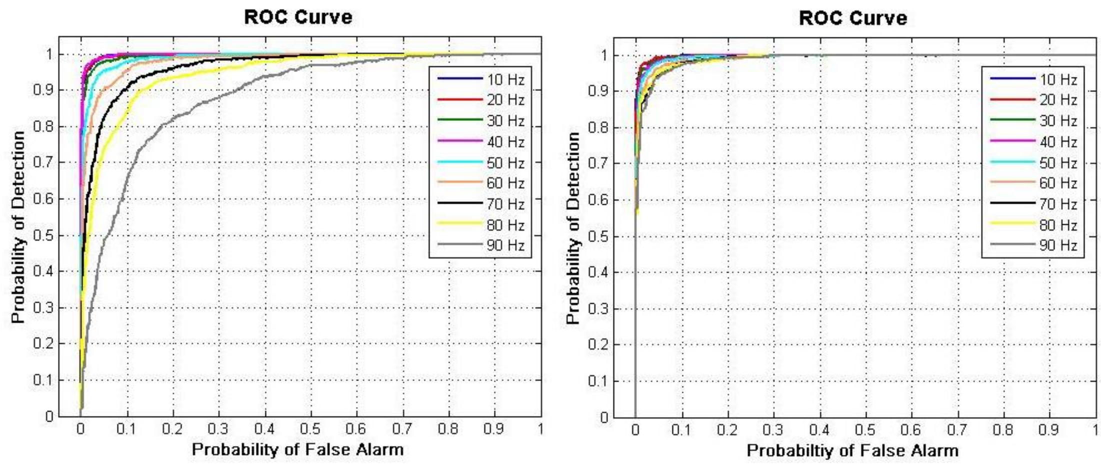

Figure 4.6: Signal detection results of varying levels of oversampling for the zero term. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz with an SNR of −10 dB. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. The zero term was thresholded to determine the presence, or absence, of a signal.

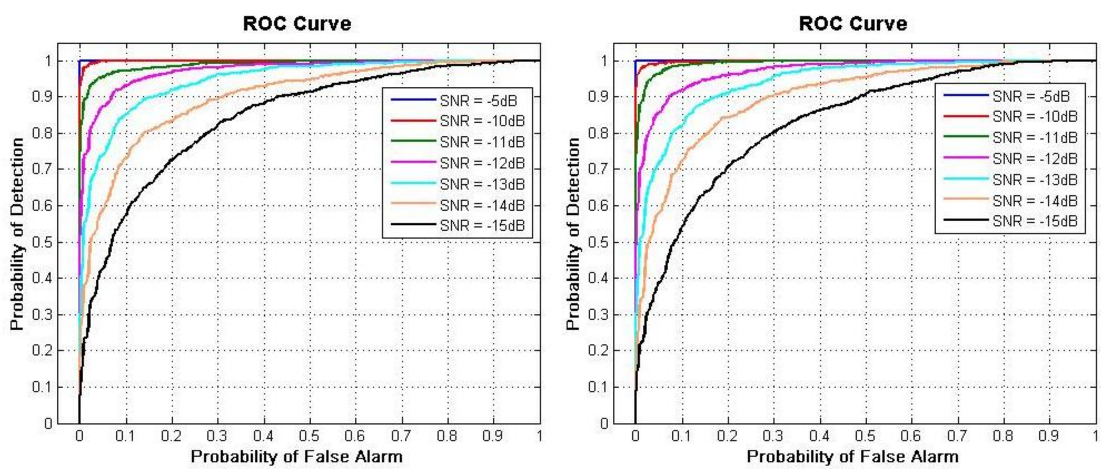

Figure 4.7: Signal detection results of varying SNR levels for the peak term. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. The peak term was thresholded to determine the presence, or absence, of a signal.

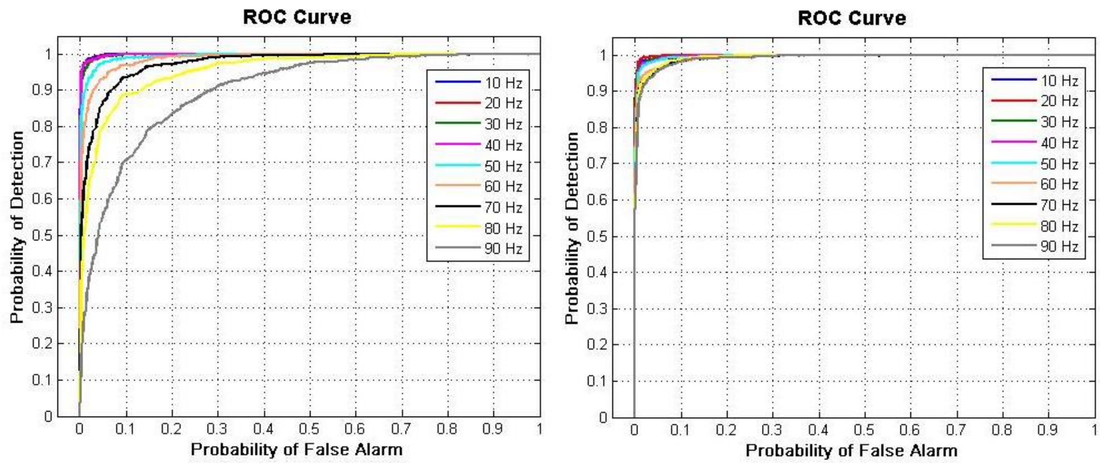

Figure 4.8: Signal detection results of varying levels of oversampling for the peak term. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz with an SNR of –10 dB. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. The peak term was thresholded to determine the presence, or absence, of a signal.

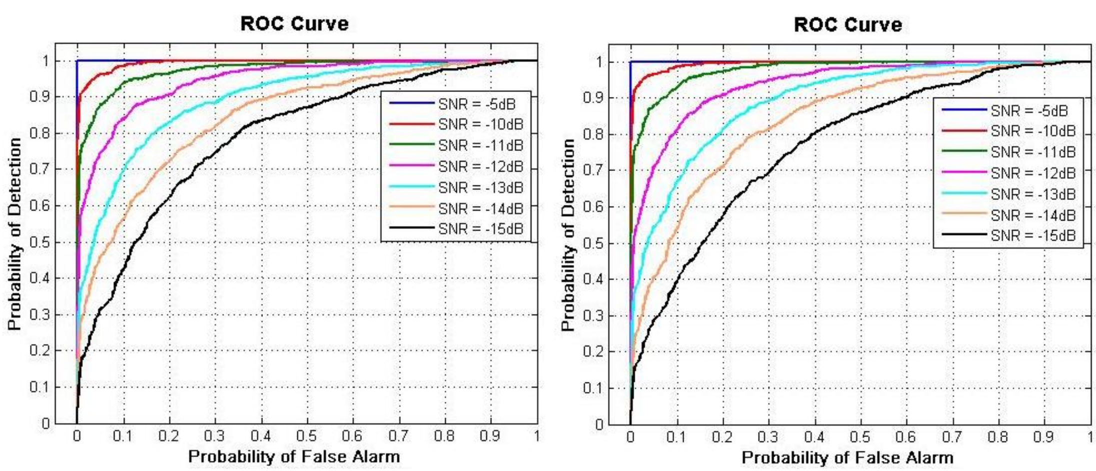

Figure 4.9: Signal detection results of varying SNR levels for the variance. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. The variance of the Intensity Analysis results was thresholded to determine the presence, or absence, of a signal.

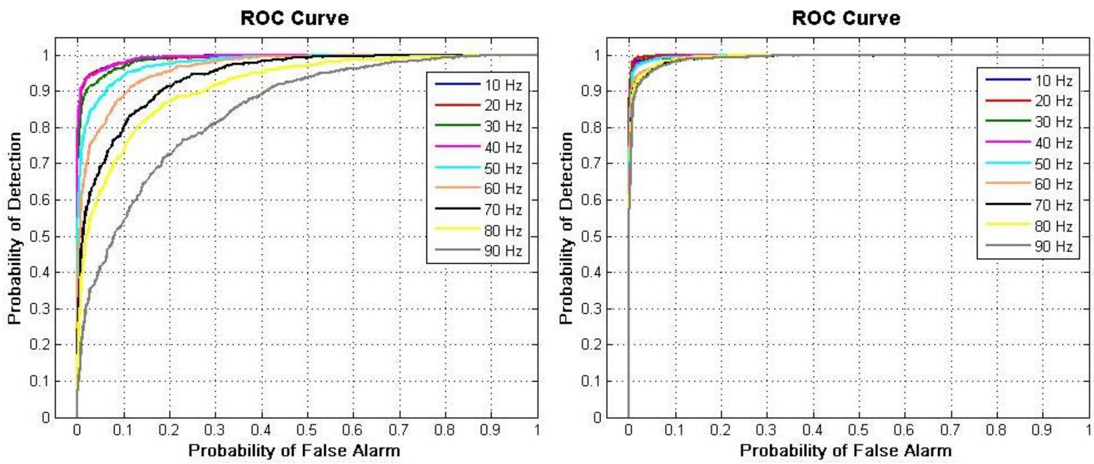

Figure 4.10: Signal detection results of varying levels of oversampling for the variance. The ROC curve on the left was created from harmonic signals and the curve on the right from PRBSs. Each signal had a 20 Hz base frequency and was sampled at 1 kHz with an SNR of −10 dB. The corresponding RPs were created with dimension of 4, delay of 12, epsilon of 1, and neighbor search using Euclidean norm between normalized vectors. The Intensity Analysis was performed with a 5 by 5 window size. The variance of the Intensity Analysis results was thresholded to determine the presence, or absence, of a signal.

5. Summary

Intensity Analysis is a new approach for the quantification of recurrence plots. It utilizes an overlapping windowing technique to give a quantified representation of the localized variations in the points of recurrence, which correspond to the presence/absence of structure within the recurrence plot.

A useful way to view the results from Intensity Analysis is to create a histogram from the results. Histograms that display larger tails and lower central regions correspond to signals constructed from relatively larger levels of deterministic signals. Conversely, histograms with lower tails and a larger central hump correspond to a greater presence of white Gaussian noise in the signal. Various metrics can be applied to the results of Intensity Analysis to determine whether or not a given recurrence plot was created from a signal containing some appreciable level of a deterministic signal.

References [12,13,14,15] have been added for additional reading on the subject of recurrence plots.

6. References

[1] P. E. Pace, “Detecting and Classifying Low Probability of Intercept Radar”, Artech House, Boston, 2004.

[2] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr., “Recurrence Quantification Analysis as an Empirical Test to Distinguish Relatively Short Deterministic Versus Random Number Series”, Physics Letters A **267**, 174-178, 2000.

[3] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr., “Recurrence Quantification Analysis and Principal Components in the Detection of Short Complex Signals”, Physics Letters A **237**, 131-135, 1998.

[4] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr., “Detecting Deterministic Signals in Exceptionally Noisy Environments Using Cross-recurrence Quantification”, Physics Letters A **246**, 122-128, 1998.

[5] L. Matassini, H. Kantz, J. Holyst, and R. Hegger, “Optimizing of Recurrence Plots for Noise Reduction”, Physical Review E **65**, 021102, 2002.

[6] J.P. Eckmann, S.O. Kamphorst and D. Ruelle, “Recurrence plots of dynamical systems,” Europhys. Lett., **4,** 973-977, 1987.

[7] F. Takens, “Detecting Strange Attractors in Turbulence”, Dynamical Systems and Turbulence (Springer Lecture Notes in Mathematics vol. 898) ed. D. Rand and L. S. Young (New York: Springer) 366-381, 1981.

[8] T. Sauer, J. A. Yorke, and M. Castagli, “Embedology”, Journal of Statistical Physics, 65(3,4) 579-616, 1991.

[9] G. P. Williams, “Chaos Theory Tamed”, Joseph Henry Press, Washington D.C., 1997.L. M. Pecora, L. Moniz, J. M. Nichols, and T. L. Carroll, “A Unified Approach to Attractor Reconstruction”, submitted to Chaos, 2006.

[10] J. S. Iwanski and E. Bradley, “Recurrence Plots of Experimental Data: To Embed or Not to Embed”, Chaos 8, 861-871, 1998.

[11] T.K. March, S.C. Chapman and R.O. Dendy, “Recurrence plot statistics and the effect of embedding,” Physica D, **200**, 171-184, 2005.

[12] Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., Kurths, J.: Recurrence Plot Based Measures of Complexity and its Application to Heart Rate Variability Data, Phys. Rev. E, **66**(2), 026702, 2002.

[13] M. Thiel, M.C. Romano, J. Kurths, R. Meucci, E. Allaria, and F.T. Arecchi, “Influence of observational noise on the recurrence quantification analysis,” Physica D, **171** 138-152, 2002.

[14] M.C. Casdagli, “Recurrence plots revisited,” Physica D **108** 12-44, 1997.

[15] J. Gao and H. Cai, “On the structures and quantification of recurrence plots,” Phys. Lett. A, **270** 75-87, 2000.